

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/662,425

Source: _____

Date Processed by STIC: _____

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:04

Input Set : N:\Crif3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

```

1 <110> APPLICANT: MINOPRIO, PAOLA
2     ARALA-CHAVES, MARIO
3     COUTINHO, ANTONIO
4     SAN MARTIN, BERNARDO REINA
5     ROUGEOT, CATHERINE
6     DEGRAVE, WIM
7     COSSON, ALAIN
8 <120> TITLE OF INVENTION: CLONING, SEQUENCING, AND EXPRESSION OF A GENE ENCODING
9     AN EUKARYOTIC AMINO ACID RACEMASE, AND DIAGNOSTIC,
10    THERAPEUTIC, AND VACCINATION APPLICATIONS OF PARASITE
11    AND VIRAL MITOGENS
12 <130> FILE REFERENCE: 03495.0200
13 <140> CURRENT APPLICATION NUMBER: US/10/662,425
14 <141> CURRENT FILING DATE: 2003-09-16
15 <150> PRIOR APPLICATION NUMBER: US/09/725,945
16 <151> PRIOR FILING DATE: 2000-11-30
17 <150> PRIOR APPLICATION NUMBER: 60/168,631
18 <151> PRIOR FILING DATE: 1999-12-03
19 <150> PRIOR APPLICATION NUMBER: 60/220,207
20 <151> PRIOR FILING DATE: 2000-07-24
21 <150> PRIOR APPLICATION NUMBER: 60/221,117
22 <151> PRIOR FILING DATE: 2000-07-27
23 <160> NUMBER OF SEQ ID NOS: 26
24 <170> SOFTWARE: PatentIn Ver. 2.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 418
28 <212> TYPE: PRT
29 <213> ORGANISM: Trypanosoma cruzi
30 <400> SEQUENCE: 1
31    Met Arg Lys Ser Val Cys Pro Lys Gln Lys Phe Phe Phe Ser Ala Phe
32         1             5             10             15
33    Pro Phe Phe Phe Phe Phe Cys Val Phe Pro Leu Ile Ser Arg Thr Gly
34         20             25             30
35    Gln Glu Lys Leu Leu Phe Asp Gln Lys Tyr Lys Ile Ile Lys Gly Glu
36         35             40             45
37    Lys Lys Glu Lys Lys Lys Asn Gln Arg Ala Asn Arg Arg Glu His Gln
38         50             55             60
39    Gln Lys Arg Glu Ile Met Arg Phe Lys Lys Ser Phe Thr Cys Ile Asp
40         65             70             75             80
41    Met His Thr Glu Gly Glu Ala Ala Arg Ile Val Thr Ser Gly Leu Pro
42         85             90             95
43    His Ile Pro Gly Ser Asn Met Ala Glu Lys Lys Ala Tyr Leu Gln Glu
44         100            105            110

```

RAW SEQUENCE LISTING

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:04

Input Set : N:\Crf3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

```

45  Asn Met Asp Tyr Leu Arg Arg Gly Ile Met Leu Glu Pro Arg Gly His
46      115      120      125
47  Asp Asp Met Phe Gly Ala Phe Leu Phe Asp Pro Ile Glu Glu Gly Ala
48      130      135      140
49  Asp Leu Gly Met Val Phe Met Asp Thr Gly Gly Tyr Leu Asn Met Cys
50      145      150      155      160
51  Gly His Asn Ser Ile Ala Ala Val Thr Ala Ala Val Glu Thr Gly Ile
52      165      170      175
53  Val Ser Val Pro Ala Lys Ala Thr Asn Val Pro Val Val Leu Asp Thr
54      180      185      190
55  Pro Ala Gly Leu Val Arg Gly Thr Ala His Leu Gln Ser Gly Thr Glu
56      195      200      205
57  Ser Glu Val Ser Asn Ala Ser Ile Ile Asn Val Pro Ser Phe Leu Tyr
58      210      215      220
59  Gln Gln Asp Val Val Val Val Leu Pro Lys Pro Tyr Gly Glu Val Arg
60      225      230      235      240
61  Val Asp Ile Ala Phe Gly Gly Asn Phe Phe Ala Ile Val Pro Ala Glu
62      245      250      255
63  Gln Leu Gly Ile Asp Ile Ser Val Gln Asn Leu Ser Arg Leu Gln Glu
64      260      265      270
65  Ala Gly Glu Leu Leu Arg Thr Glu Ile Asn Arg Ser Val Lys Val Gln
66      275      280      285
67  His Pro Gln Leu Pro His Ile Asn Thr Val Asp Cys Val Glu Ile Tyr
68      290      295      300
69  Gly Pro Pro Thr Asn Pro Glu Ala Asn Tyr Lys Asn Val Val Ile Phe
70      305      310      315      320
71  Gly Asn Arg Gln Ala Asp Arg Gly Thr Ser Ala Lys Met Ala Thr Leu
72      325      330      335
73  Tyr Ala Lys Gly Gln Leu Arg Ile Gly Glu Thr Phe Val Tyr Glu Ser
74      340      345      350
75  Ile Leu Gly Ser Leu Phe Gln Gly Arg Val Leu Gly Glu Arg Ile
76      355      360      365
77  Pro Gly Val Lys Val Pro Val Thr Lys Asp Ala Glu Glu Gly Met Leu
78      370      375      380
79  Val Val Thr Ala Glu Ile Thr Gly Lys Ala Phe Ile Met Gly Phe Asn
80      385      390      395      400
81  Thr Met Leu Phe Asp Pro Thr Asp Pro Phe Lys Asn Gly Phe Thr Leu
82      405      410      415
83  Lys Gln
85 <210> SEQ ID NO: 2
86 <211> LENGTH: 389
87 <212> TYPE: PRT
88 <213> ORGANISM: Trypanosoma cruzi
89 <400> SEQUENCE: 2
90  Arg Thr Gly Gln Glu Lys Leu Leu Phe Asp Gln Lys Tyr Lys Ile Ile
91      1      5      10      15
92  Lys Gly Glu Lys Lys Glu Lys Lys Lys Asn Gln Arg Ala Asn Arg Arg
93      20      25      30
94  Glu His Gln Gln Lys Arg Glu Ile Met Arg Phe Lys Lys Ser Phe Thr

```

RAW SEQUENCE LISTING

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:04

Input Set : N:\Crif3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

```

95          35          40          45
96 Cys Ile Asp Met His Thr Glu Gly Glu Ala Ala Arg Ile Val Thr Ser
97          50          55          60
98 Gly Leu Pro His Ile Pro Gly Ser Asn Met Ala Glu Lys Lys Ala Tyr
99          65          70          75          80
100 Leu Gln Glu Asn Met Asp Tyr Leu Arg Arg Gly Ile Met Leu Glu Pro
101          85          90          95
102 Arg Gly His Asp Asp Met Phe Gly Ala Phe Leu Phe Asp Pro Ile Glu
103          100          105          110
104 Glu Gly Ala Asp Leu Gly Met Val Phe Met Asp Thr Gly Gly Tyr Leu
105          115          120          125
106 Asn Met Cys Gly His Asn Ser Ile Ala Ala Val Thr Ala Ala Val Glu
107          130          135          140
108 Thr Gly Ile Val Ser Val Pro Ala Lys Ala Thr Asn Val Pro Val Val
109          145          150          155          160
110 Leu Asp Thr Pro Ala Gly Leu Val Arg Gly Thr Ala His Leu Gln Ser
111          165          170          175
112 Gly Thr Glu Ser Glu Val Ser Asn Ala Ser Ile Ile Asn Val Pro Ser
113          180          185          190
114 Phe Leu Tyr Gln Gln Asp Val Val Val Val Leu Pro Lys Pro Tyr Gly
115          195          200          205
116 Glu Val Arg Val Asp Ile Ala Phe Gly Gly Asn Phe Phe Ala Ile Val
117          210          215          220
118 Pro Ala Glu Gln Leu Gly Ile Asp Ile Ser Val Gln Asn Leu Ser Arg
119          225          230          235          240
120 Leu Gln Glu Ala Gly Glu Leu Leu Arg Thr Glu Ile Asn Arg Ser Val
121          245          250          255
122 Lys Val Gln His Pro Gln Leu Pro His Ile Asn Thr Val Asp Cys Val
123          260          265          270
124 Glu Ile Tyr Gly Pro Pro Thr Asn Pro Glu Ala Asn Tyr Lys Asn Val
125          275          280          285
126 Val Ile Phe Gly Asn Arg Gln Ala Asp Arg Gly Thr Ser Ala Lys Met
127          290          295          300
128 Ala Thr Leu Tyr Ala Lys Gly Gln Leu Arg Ile Gly Glu Thr Phe Val
129          305          310          315          320
130 Tyr Glu Ser Ile Leu Gly Ser Leu Phe Gln Gly Arg Val Leu Gly Glu
131          325          330          335
132 Glu Arg Ile Pro Gly Val Lys Val Pro Val Thr Lys Asp Ala Glu Glu
133          340          345          350
134 Gly Met Leu Val Val Thr Ala Glu Ile Thr Gly Lys Ala Phe Ile Met
135          355          360          365
136 Gly Phe Asn Thr Met Leu Phe Asp Pro Thr Asp Pro Phe Lys Asn Gly
137          370          375          380
138 Phe Thr Leu Lys Gln
139          385
141 <210> SEQ ID NO: 3
142 <211> LENGTH: 29
143 <212> TYPE: PRT
144 <213> ORGANISM: Trypanosoma cruzi

```

RAW SEQUENCE LISTING

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:04

Input Set : N:\CrF3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

```

145 <400> SEQUENCE: 3
146   Met Arg Lys Ser Val Cys Pro Lys Gln Lys Phe Phe Phe Ser Ala Phe
147       1               5               10               15
148   Pro Phe Phe Phe Phe Phe Cys Val Phe Pro Leu Ile Ser
149               20               25
151 <210> SEQ ID NO: 4
152 <211> LENGTH: 354
153 <212> TYPE: PRT
154 <213> ORGANISM: Trypanosoma cruzi
155 <400> SEQUENCE: 4
156   Met Arg Phe Lys Lys Ser Phe Thr Cys Ile Asp Met His Thr Glu Gly
157       1               5               10               15
158   Glu Ala Ala Arg Ile Val Thr Ser Gly Leu Pro His Ile Pro Gly Ser
159               20               25               30
160   Asn Met Ala Glu Lys Lys Ala Tyr Leu Gln Glu Asn Met Asp Tyr Leu
161               35               40               45
162   Arg Arg Gly Ile Met Leu Glu Pro Arg Gly His Asp Asp Met Phe Gly
163               50               55               60
164   Ala Phe Leu Phe Asp Pro Ile Glu Glu Gly Ala Asp Leu Gly Met Val
165               65               70               75               80
166   Phe Met Asp Thr Gly Gly Tyr Leu Asn Met Cys Gly His Asn Ser Ile
167               85               90               95
168   Ala Ala Val Thr Ala Ala Val Glu Thr Gly Ile Val Ser Val Pro Ala
169               100              105              110
170   Lys Ala Thr Asn Val Pro Val Val Leu Asp Thr Pro Ala Gly Leu Val
171               115              120              125
172   Arg Gly Thr Ala His Leu Gln Ser Gly Thr Glu Ser Glu Val Ser Asn
173               130              135              140
174   Ala Ser Ile Ile Asn Val Pro Ser Phe Leu Tyr Gln Gln Asp Val Val
175               145              150              155              160
176   Val Val Leu Pro Lys Pro Tyr Gly Glu Val Arg Val Asp Ile Ala Phe
177               165              170              175
178   Gly Gly Asn Phe Phe Ala Ile Val Pro Ala Glu Gln Leu Gly Ile Asp
179               180              185              190
180   Ile Ser Val Gln Asn Leu Ser Arg Leu Gln Glu Ala Gly Glu Leu Leu
181               195              200              205
182   Arg Thr Glu Ile Asn Arg Ser Val Lys Val Gln His Pro Gln Leu Pro
183               210              215              220
184   His Ile Asn Thr Val Asp Cys Val Glu Ile Tyr Gly Pro Pro Thr Asn
185               225              230              235              240
186   Pro Glu Ala Asn Tyr Lys Asn Val Val Ile Phe Gly Asn Arg Gln Ala
187               245              250              255
188   Asp Arg Ser Pro Cys Gly Thr Gly Thr Ser Ala Lys Met Ala Thr Leu
189               260              265              270
190   Tyr Ala Lys Gly Gln Leu Arg Ile Gly Glu Thr Phe Val Tyr Glu Ser
191               275              280              285
192   Ile Leu Gly Ser Leu Phe Gln Gly Arg Val Leu Gly Glu Glu Arg Ile
193               290              295              300
194   Pro Gly Val Lys Val Pro Val Thr Lys Asp Ala Glu Glu Gly Met Leu

```

RAW SEQUENCE LISTING

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:04

Input Set : N:\Crf3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

```

195      305      310      315      320
196 Val Val Thr Ala Glu Ile Thr Gly Lys Ala Phe Ile Met Gly Phe Asn
197      325      330      335
198 Thr Met Leu Phe Asp Pro Thr Asp Pro Phe Lys Asn Gly Phe Thr Leu
199      340      345      350
200 Lys Gln
202 <210> SEQ ID NO: 5
203 <211> LENGTH: 330
204 <212> TYPE: PRT
205 <213> ORGANISM: Clostridium sticklandii
206 <400> SEQUENCE: 5
207 Met Lys Phe Ser Lys Gly Ile His Ala Ile Asp Ser His Thr Met Gly
208 1 5 10 15
209 Glu Pro Thr Arg Ile Val Val Gly Gly Ile Pro Gln Ile Asn Gly Glu
210 20 25 30
211 Thr Met Ala Asp Lys Lys Lys Tyr Leu Glu Asp Asn Leu Asp Tyr Val
212 35 40 45
213 Arg Thr Ala Leu Met His Glu Pro Arg Gly His Asn Asp Met Phe Gly
214 50 55 60
215 Ser Ile Ile Thr Ser Ser Asn Asn Lys Glu Ala Asp Phe Gly Ile Ile
216 65 70 75 80
217 Phe Met Asp Gly Gly Gly Tyr Leu Asn Met Cys Gly His Gly Ser Ile
218 85 90 95
219 Gly Ala Ala Thr Val Ala Val Glu Thr Gly Met Val Glu Met Val Glu
220 100 105 110
221 Pro Val Thr Asn Ile Asn Met Glu Ala Pro Ala Gly Leu Ile Lys Ala
222 115 120 125
223 Lys Val Met Val Glu Asn Glu Lys Val Lys Glu Val Ser Ile Thr Asn
224 130 135 140
225 Val Pro Ser Phe Leu Tyr Met Glu Asp Ala Lys Leu Glu Val Pro Ser
226 145 150 155 160
227 Leu Asn Lys Thr Ile Thr Phe Asp Ile Ser Phe Gly Gly Ser Phe Phe
228 165 170 175
229 Ala Ile Ile His Ala Lys Glu Leu Gly Val Lys Val Glu Thr Ser Gln
230 180 185 190
231 Val Asp Val Leu Lys Lys Leu Gly Ile Glu Ile Arg Asp Leu Ile Asn
232 195 200 205
233 Glu Lys Ile Lys Val Gln His Pro Glu Leu Glu His Ile Lys Thr Val
234 210 215 220
235 Asp Leu Val Glu Ile Tyr Asp Glu Pro Ser Asn Pro Glu Ala Thr Tyr
236 225 230 235 240
237 Lys Asn Val Val Ile Phe Gly Gln Gly Gln Val Asp Arg Gly Thr Ser
238 245 250 255
239 Ala Lys Leu Ala Thr Leu Tyr Lys Lys Gly His Leu Lys Ile Asp Glu
240 260 265 270
241 Lys Glu Val Tyr Glu Ser Ile Thr Gly Thr Met Phe Lys Gly Arg Val
242 275 280 285
243 Leu Glu Glu Thr Lys Val Gly Glu Phe Asp Ala Ile Ile Pro Glu Ile
244 290 295 300

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/662,425

DATE: 10/20/2004
TIME: 15:11:05

Input Set : N:\Crf3\RULE60\10662425.raw.txt
Output Set: N:\CRF4\10202004\J662425.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:12; N Pos. 3,12,15
Seq#:13; N Pos. 6,12,15
Seq#:14; N Pos. 3,12,15

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 8

VERIFICATION SUMMARY

DATE: 10/20/2004

PATENT APPLICATION: US/10/662,425

TIME: 15:11:05

Input Set : N:\Crf3\RULE60\10662425.raw.txt

Output Set: N:\CRF4\10202004\J662425.raw

L:439 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:442 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:12
L:443 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:451 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:454 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:13
L:455 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:463 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:466 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:14
L:467 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0